Int J CARS (2016) 11:1673–1685
DOI 10.1007/s11548-016-1400-9

ORIGINAL ARTICLE

Fully automatic reconstruction of personalized 3D volumes of the proximal femur from 2D X-ray images

Weimin Yu[1] · **Chengwen Chu**[1] · **Moritz Tannast**[2] · **Guoyan Zheng**[1]

Received: 1 September 2015 / Accepted: 21 March 2016 / Published online: 2 April 2016
© CARS 2016

Abstract
Purpose Accurate preoperative planning is crucial for the outcome of total hip arthroplasty. Recently, 2D pelvic X-ray radiographs have been replaced by 3D CT. However, CT suffers from relatively high radiation dosage and cost. An alternative is to reconstruct a 3D patient-specific volume data from 2D X-ray images.
Methods In this paper, based on a fully automatic image segmentation algorithm, we propose a new control point-based 2D–3D registration approach for a deformable registration of a 3D volumetric template to a limited number of 2D calibrated X-ray images and show its application to personalized reconstruction of 3D volumes of the proximal femur. The 2D–3D registration is done with a hierarchical two-stage strategy: the scaled-rigid 2D–3D registration stage followed by a regularized deformable B-spline 2D–3D registration stage. In both stages, a set of control points with uniform spacing are placed over the domain of the 3D volumetric template first. The registration is then driven by computing updated positions of these control points with intensity-based 2D–2D image registrations of the input X-ray images with the associated digitally reconstructed radiographs, which allows computing the associated registration transformation at each stage.
Results Evaluated on datasets of 44 patients, our method achieved an overall surface reconstruction accuracy of 0.9 ± 0.2 mm and an average Dice coefficient of 94.4 ± 1.1 %. We further investigated the cortical bone region reconstruction accuracy, which is important for planning cementless total hip arthroplasty. An average cortical bone region Dice coefficient of 85.1 ± 2.9 % and an inner cortical bone surface reconstruction accuracy of 0.7 ± 0.2 mm were found.
Conclusions In summary, we developed a new approach for reconstruction of 3D personalized volumes of the proximal femur from 2D X-ray images. Comprehensive experiments demonstrated the efficacy of the present approach.

Keywords Segmentation · Deformable registration · Reconstruction · 2D–3D · Random forest regression · B-spline

Introduction

Accurate and reliable preoperative templating is essential in contemporary total hip arthroplasty (THA) as it aids the surgeon in the selection of appropriate implant geometry, size, and position [1]. This is especially true for a cementless THA as the endosteal fit of the femoral component is a major factor determining load transfer and, consequently, periprosthetic bone remodeling. Several studies have shown that femoral fit predicts not only radiographic changes following THA [2], but also affects clinical results in terms of implant survival [3]. Although two-dimensional (2D) plain radiographs are the standard imaging means that are widely used in planning THA, several studies have demonstrated that radiographic assessment of femoral canal shape and geometry has limited reliability [4,5]. A recent development was the introduction

Electronic supplementary material The online version of this article (doi:10.1007/s11548-016-1400-9) contains supplementary material, which is available to authorized users.

✉ Guoyan Zheng
guoyan.zheng@ieee.org; guoyan.zheng@istb.unibe.ch

[1] Institute for Surgical Technology and Biomechanics, University of Bern, Stauffacherstr. 78, Bern 3014, Switzerland

[2] Department of Orthopaedic Surgery, Inselspital, University of Bern, Bern, Switzerland

of computed tomography (CT) for three-dimensional (3D) THA planning.

In comparison with plain radiograph, CT-based 3D planning of THA offers several advantages [1,6]. More specifically, CT has the benefits of avoiding errors resulting from magnification and inaccurate patient positioning. Additional benefits include the assessment in the axial plane, replacement of 2D projections with 3D data, and the availability of information on bone quality including accurate differentiation between cortical and cancellous bone. The concern on 3D CT-based planning of THA, however, lies in the increase in radiation dosage to the patients [6,7]. Even though it was claimed in [6] that the increase could be justified, the dose of a preoperative CT for THA was increased by at least 30 % compared to conventional radiographs. An alternative is to reconstruct a patient-specific 3D volume data from 2D X-rays.

Depending on the output, 2D–3D reconstruction methods can be largely classified into two categories: 3D surface model reconstruction methods [8–10] and 3D volume reconstruction methods [11–14]. The methods in the former category compute 3D patient-specific surface models from one or multiple 2D X-ray images. No intensity information or information about cortical bone is available. The reconstructed surface models may be used to plan cup component in THA as shown in [15], but not for planning cementless stem component due to the missing information about proximal femur morphology and the intramedullary anatomy.

The methods in the second category generate 3D patient-specific volumes from a limited number of X-ray images. When two or more C-arm/X-ray images are available, Yao and Tayor [13] and Sadowsky et al. [12] proposed an iterative registration process to estimate the pose, scale, and modes of variation in a tetrahedral mesh-based Statistical Shape and Intensity Model (SSIM) by minimizing the difference between the simulated digitally reconstructed radiographs (DRRs) and the real X-ray images. Mutual information was used as the similarity measure. With leave-one-out tests, an average registration error of 2.0 mm was reported in [12]. Zheng [14] proposed to reconstruct a patient-specific 3D volume by matching independent shape and appearance models that are learned from a set of training data to a limited number of C-arm/X-ray images. An intensity-based nonrigid 2D–3D registration algorithm was proposed to deformably fit the learned models to the input images. When two C-arm images were used, a mean reconstruction accuracy of 1.5 mm was reported in [14].

In this paper, we propose a new control-point-based 2D–3D registration approach for a deformable registration of a 3D volumetric template to a limited number of 2D X-ray images and show its application to a personalized reconstruction of volumes of the proximal femur. In comparison with previous work, our contributions are as follows:

- We propose a novel control-point-based 2D–3D nonrigid registration method for a personalized reconstruction of a 3D volume of the proximal femur from 2D X-ray images.
- Our method is fully automatic due to the application of a previously developed machine-learning-based technique [16] for contour extraction and landmark-based initialization.
- We conduct comprehensive experiments for a thorough investigation of our approach.
- We evaluate not only the volume reconstruction accuracy but also the cortical bone region reconstruction accuracy, which is important for stem planning in cementless THA.

Materials and methods

Overview

Figure 1 shows an overview of the complete workflow of our method. Here, without loss of generality, we assume that the given pair of X-ray images, one acquired along the antero-posterior (AP) direction and the other along the axial (AX) direction, has been calibrated and registered with a common reference space c. As we would like to match a 3D volumetric template to the 2D calibrated X-ray images to derive a 3D patient-specific volume, we consider the 3D volumetric template as the floating image $\{\mathbf{I}(\mathbf{x_f})\}$, where $\mathbf{x_f}$ is a point in the template volume, and the 2D is calibrated X-ray images as the reference images. The template is aligned to the X-ray reference space $\mathbf{c}$ by following forward mapping:

$$\mathbf{I}(\mathbf{x_c}\left(T_g, T_l\right)) = \mathbf{I}\left(T_g \cdot T_l \cdot \mathbf{x_f}\right) \tag{1}$$

where T_g is a similarity transformation and T_l is a local deformation.

In order to compute T_g and T_l, our method starts with X-ray image segmentation and template initialization. The extracted femoral regions and the initialized volumetric template are then used for an iterative control-point-based 2D–3D registration as detailed below.

X-ray image segmentation and template volume initialization

Femoral contour in any one of the given pair of X-ray images is fully automatically segmented using the approach that we introduced in [16] (Fig. 2). In our method, a contour is described by a set of ordered landmarks. We have chosen to use $L_{AP} = 57$ landmarks to express a contour in an AP image and $L_{LM} = 54$ landmarks to describe a contour in an AX image (see Fig. 3a, for details about how the landmarks are ordered in different views). Figure 3b, shows

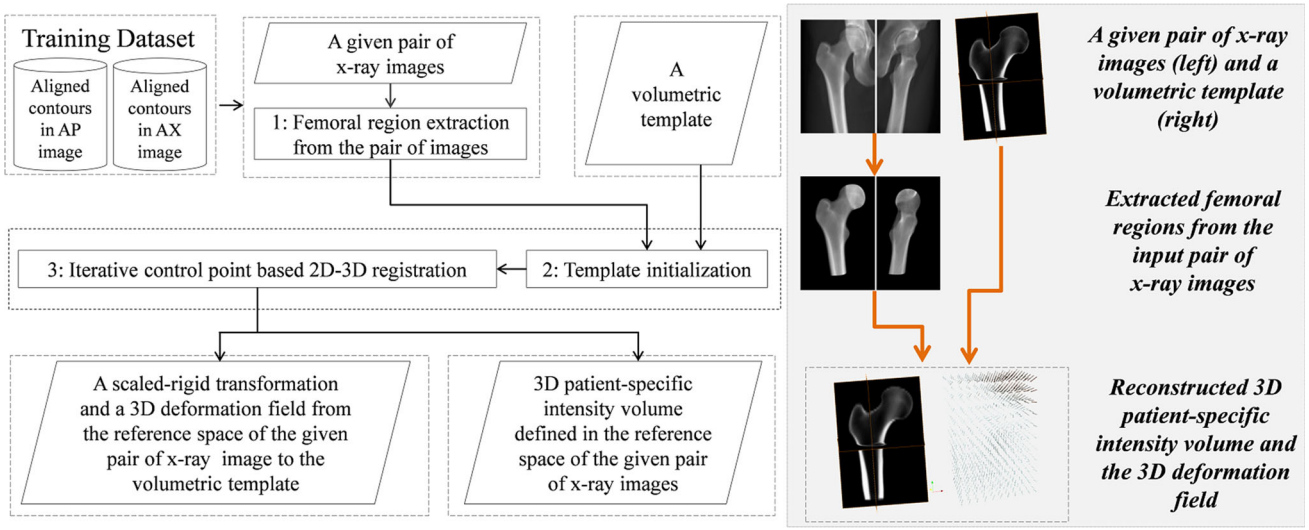

Fig. 1 Flowchart of the proposed 2D–3D volume reconstruction method

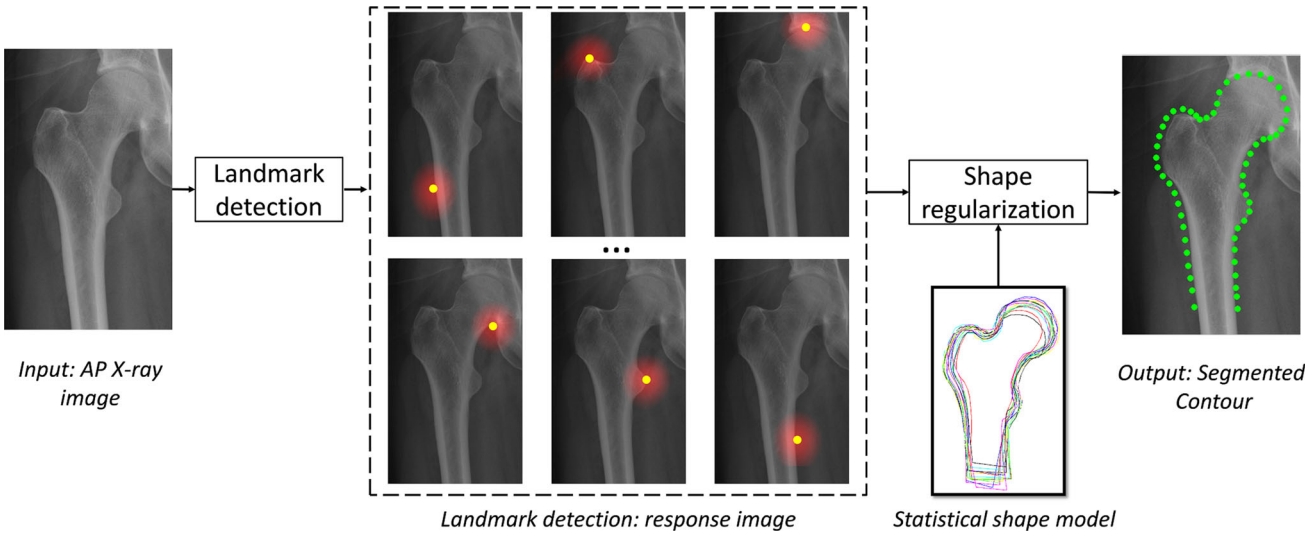

Fig. 2 A schematic illustration of the pipeline for the contour extraction in a given AP image. See [16] for details

several examples of the extracted femoral contours in AP and AX views, respectively. The extracted contours in each image are used to extract the region of interest (ROI) of the proximal femur, which is done by setting the value for any pixel outside the associated ROI to 0 (see Fig. 4 for examples).

From the extracted contours, we can use part of the ordered landmarks as shown in Fig. 3b, to define projections of three anatomical landmarks, i.e., the greater trochanter (GT), the lesser trochanter (LT), and the center of the femoral head (FH). More specifically, from the contour extracted from an AP image, the projection of the greater trochanter and the projection of the lesser trochanter are defined by landmarks 20 and 48, respectively. To estimate the projection of the femoral head, we choose to do a circle fitting using landmarks indexed in the range from 27 to 41 and the center of the fitted circle is regarded as the projection of the femoral head. Similarly, we can find projections of these three landmarks in an AX image. From the corresponding projections of an anatomical landmark in the given pair of images, we can use a triangulation-based approach to calculate the position of the landmark. Let's denote the positions of the three anatomical landmarks in the X-ray reference space $\mathbf{c}$ as $\mathrm{GT}_{\mathbf{c}}$, $\mathrm{LT}_{\mathbf{c}}$, $\mathrm{FH}_{\mathbf{c}}$, respectively.

Accordingly, we can also obtain the positions of the same three landmarks in the space of the volumetric template.

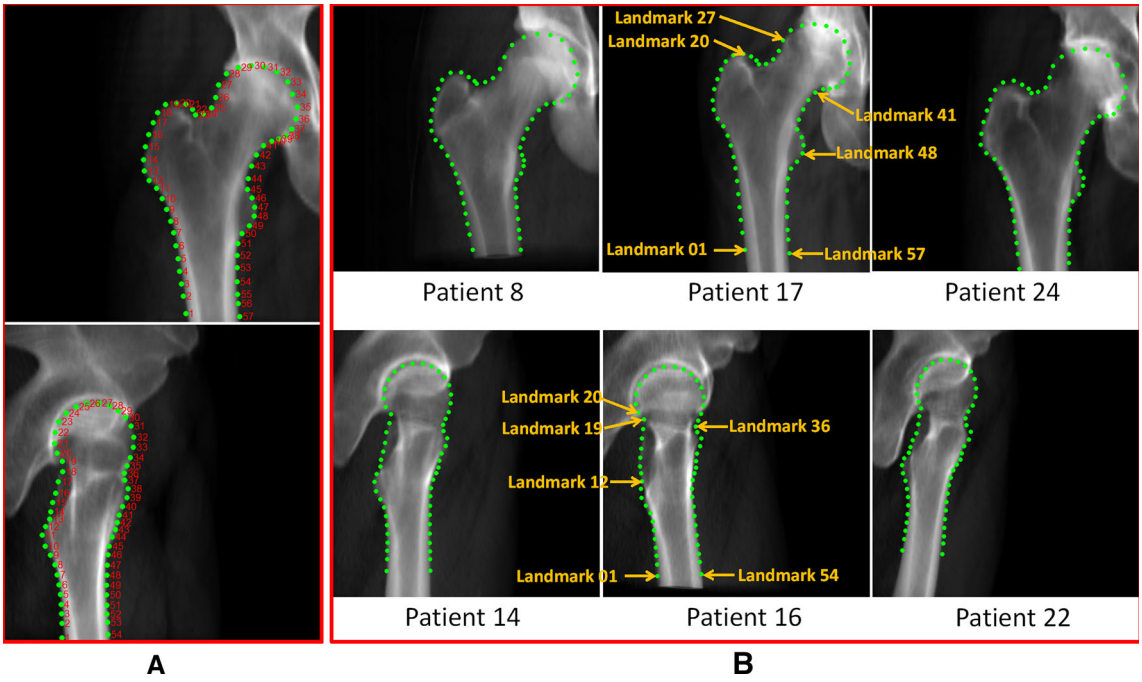

Fig. 3 **a** Ordered landmarks for defining the femoral contours in an image acquired along the AP direction (*top*, 57 ordered landmarks) and in an image acquired along the AX direction (*bottom*, 54 ordered landmarks), and **b** Examples of fully automatic detection of contours from AP DRRs (*top*) and from AX DRRs (*bottom*). Indexes for defining important anatomical landmarks are shown

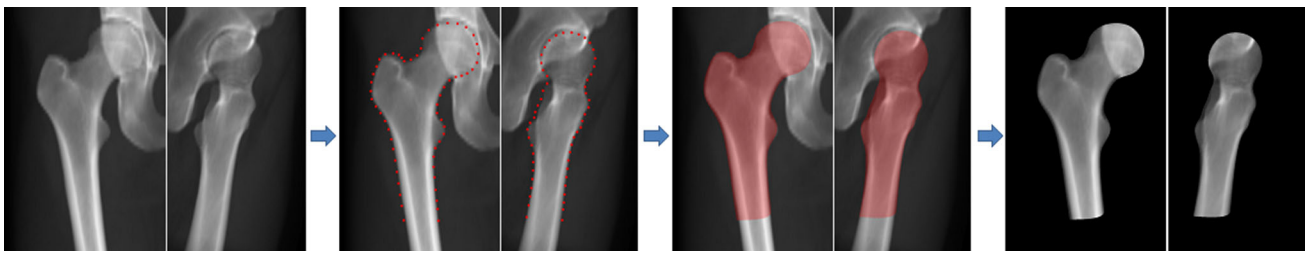

Fig. 4 Femoral ROI extraction. Please note that all images are cropped to save space. *Left* a given pair of X-ray images; *middle-left* after automatic contour extraction; *middle right* after binary ROI mask generation; *right* the extract femoral regions

We denote the three points in the space of the volumetric template as $\mathrm{GT}_{\mathbf{f}}$, $\mathrm{LT}_{\mathbf{f}}$, $\mathrm{FH}_{\mathbf{f}}$, respectively. Thus, given the corresponding coordinates of these three anatomical landmarks, we can estimate an initial similarity transformation T_g^0 from the space of the volumetric template to the reference space of the given pair of X-ray images.

Control-point-based 2D–3D registration

Problem formulation

Given the extracted femoral ROIs and the initial estimation of the similarity transformation, our goal at this step is to estimate T_g and T_l as shown in Eq. (1) which should be interpreted as follows. Given a voxel $\mathbf{x_f}$ in the template volume space, the destination of this voxel under the forward transformation is $\mathbf{x_c}\left(T_g, T_l\right) = T_g \cdot T_l \cdot \mathbf{x_f}$. The aligned volume at voxel $\mathbf{x_c}\left(T_g, T_l\right)$ is set to the intensity $\mathbf{I}\left(\mathbf{x_f}\right)$, which then allows creating DRRs by simulating X-ray projections. To avoid holes in the aligned volume, a backward warping as follows is used.

$$\mathbf{x_f}\left(T_g, T_l\right) = \left(T_l\right)^{-1} \cdot \left(T_g\right)^{-1} \cdot \mathbf{x_c} \tag{2}$$

It is straightforward to compute the inverse of the similarity transformation T_g. However, it is time-consuming to compute the inverse of the forward local deformation T_l. In [14], the inversion of the forward local deformation is done by using a fixed-point approach as introduced in [17].

In this paper, we solve the problem differently. The nonrigid registration is done with a hierarchical two-stage strategy as shown in Fig. 5: a scaled-rigid 2D–3D registration

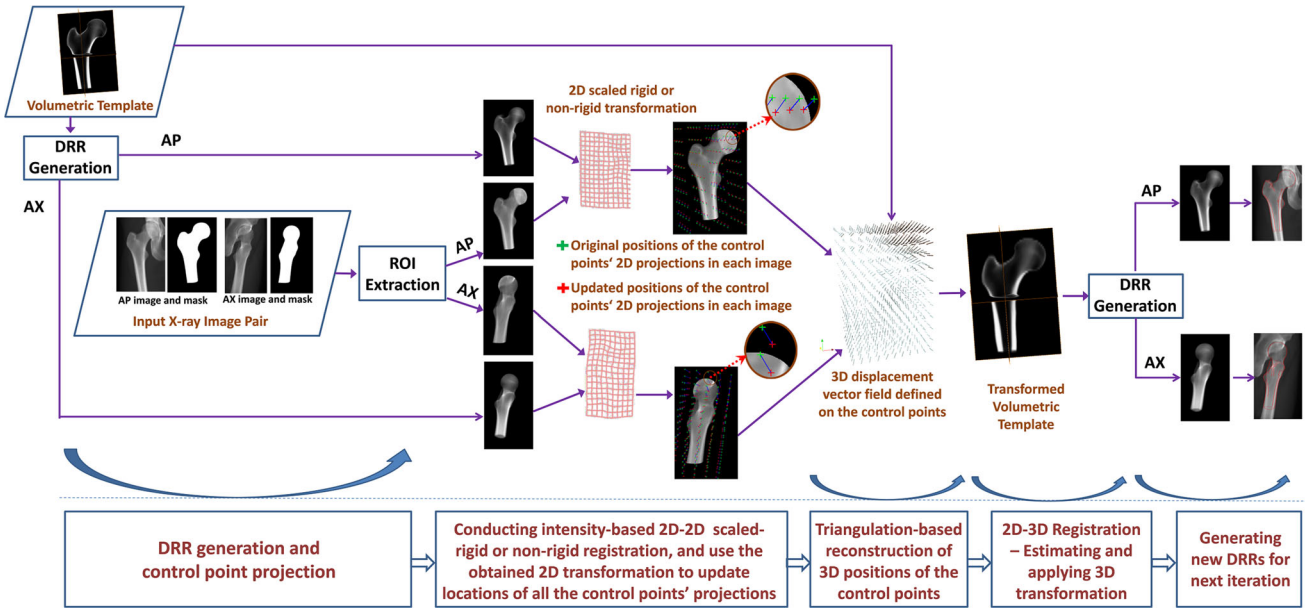

Fig. 5 A schematic illustration of the control-point-based 2D–3D registration process. See text for details

stage followed by a regularized deformable B-spline 2D–3D registration stage. In the first stage, the similarity transformation T_g is estimated, while in the second stage, we directly estimate $(T_l)^{-1}$, which is a deformation field from the X-ray reference space to the floating template space. Obtaining this deformation field will allow one to warp the floating volumetric template to the X-ray reference space. No inversion of the forward deformation field is required. Details about these two stages will be presented below.

Registration method

In both stages, a set of 3D control points with uniform spacing are placed over the domain of the volumetric template as follows (see Fig. 6 for a schematic illustration in 2D). Denote the domain of the volumetric template as $\mathbf{\Omega} = \{(x, y, z)\,|0 \leq x \leq n_1, 0 \leq y \leq n_2, 0 \leq z \leq n_3\}$, where n_1, n_2, n_3 denote the number of voxels along x-, y-, and z-coordinate directions, respectively. Choosing three positive integers $s_1, s_2, s_3 \geq 1$, describing the spacing between any two control points along each coordinate, we can define a mesh G of the control points by setting:

$$G := s_1\mathbb{Z}_{z_1} \times s_2\mathbb{Z}_{z_2} \times s_3\mathbb{Z}_{z_3} \tag{3}$$

with $z_i := \frac{n_i}{s_i} + 1 \in \mathbb{N}$ and $\mathbb{Z}_{z_i} := \{k \in \mathbb{Z} | -1 \leq k \leq z_i + 1\}$.

Such setting guarantees that every voxel of the volumetric template is surrounded by a local mesh cube of size $4 \times 4 \times 4$ control points due to the fact that a control point is placed in every corner of $\mathbf{\Omega}$ and an extra layer of control points around $\mathbf{\Omega}$ is added. There are in total $(z_1 + 3) \times (z_2 + 3) \times (z_3 + 3)$ control points. Below, we present the details about the two registration stages.

Scaled-rigid 2D–3D registration stage (stage-scaledRigid2D3D)

Given an initial estimation of the similarity transformation T_g^0, the control-point-based scaled-rigid 2D–3D registration will be executed as follows:

- *Step 1: DRR Generation and Control Point Projection.* Based on an estimation of the scaled-rigid transformation T_g^t obtained at the tth iteration, we generate DRRs using Nvidia's CUDA environment [18]. At the same time, using T_g^t, we transform all control points from the floating volume space to the X-ray reference space. We denote an arbitrary control point with index i, j, k as $L_{ijk}^{sr,t}$. After that, we do a forward projection of all transformed control point to the input X-ray images.
- *Step 2: 2D–2D Intensity-based Image Registration.* At this step, an intensity-based scaled-rigid 2D–2D registration between each 2D DRR and its associated 2D X-ray image is performed using the intensity-based registration toolbox "elastix" [19] to estimate a 2D similarity transform, which is then used to update the localizations of the 2D projections of all control points. For the intensity-based scaled-rigid 2D–2D registration, we choose to use Mattes mutual information [20] as the similarity metric and the adaptive stochastic gradient descent optimization [21] as the optimization method.

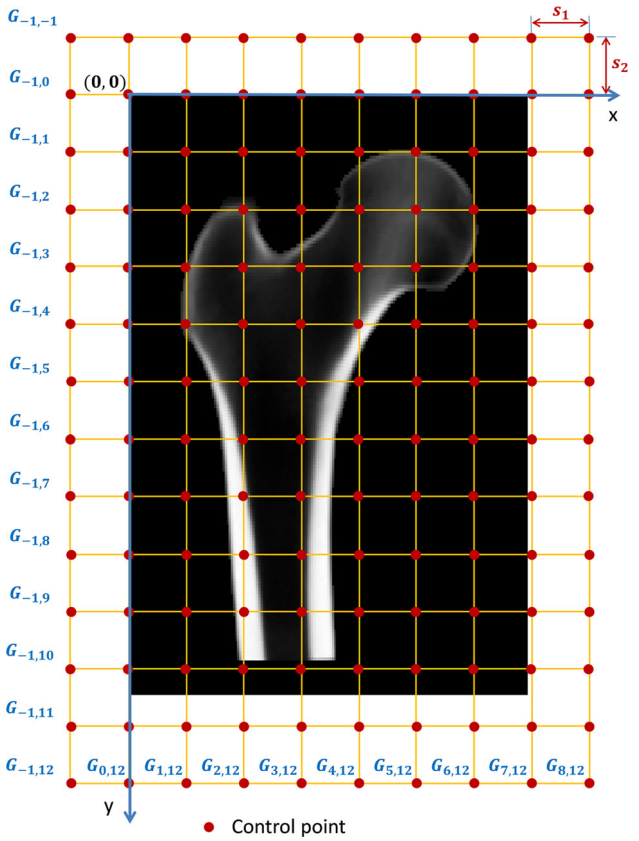

Fig. 6 A schematic illustration of how to place control point grid around a 2D image template

- *Step 3: Triangulation-based Point Reconstruction*. Given the updated 2D locations of the projections of a 3D control point with index i, j, k in the input images, an updated 3D position $L_{ijk}^{sr,t+1}$ of the control point $L_{ijk}^{sr,t}$ can be reconstructed from those updated 2D locations via a triangulation strategy as shown in Fig. 7.
- *Step 4: 2D–3D Registration*. Given two sets of 3D positions $\{L_{ijk}^{sr,t}\}$ and $\{L_{ijk}^{sr,t+1}\}$ with known correspondences, we can estimate an updated 3D similarity transformation T_g^{t+1} [22].

Repeat **Steps 1–4** a few times (we empirically chose two times) to get an estimation of the 3D similarity transformation T_g.

Regularized deformable B-spline 2D–3D registration stage (stage-nonRigid2D3D)

The estimated transformation T_g allows one to align the floating volume to the X-ray reference space. After that, the control point-based deformable B-spline 2D–3D registration stage will be executed.

- *Step 1: DRR Generation and Control Point Projection*. At this step, we first generate DRRs from a warped volume $\{\mathbf{I}_\mathbf{c}^t\}$ obtained at the tth iteration. We then use T_g to transform all control points from the floating volume space to the X-ray reference space and do a forward projection of transformed control point. We denote a transformed control point with index i, j, k as $L_{ijk}^{nr,t}$.
- *Step 2: 2D–2D Intensity-based Image Registration*. At this step, we conduct an intensity-based deformable B-spline 2D–2D registration of each X-ray image with the associated DRR by minimizing a sequential combination of Mattes mutual information [20] with residual complexity [23]. The reason why we choose to do this is because empirically we find that the residual complexity helps to achieve accurate results, but it has a small capture range, while the Mattes mutual information has relatively larger capture range, but it is less accurate. The obtained 2D deformation field will be used to update the locations of the 2D projections of all control points. For details about how to compute Mattes mutual information, we refer to [20]. Below, we briefly present how to compute residual complexity. For details, we refer to [23].
  Residual complexity was originally introduced by Myronenko and Song [23] as an efficient image similarity measure for handling intensity nonstationarities and complex spatially varying intensity distortions in monomodal settings. In this study, after a deformable B-spline 2D–2D registration of each X-ray image with the associated DRR by minimizing Mattes mutual information [20], we further conduct a deformable B-spline 2D–2D registration by minimizing residual complexity. Assuming that the qth X-ray image is $\mathbf{I}_{X-ray}^q$ and the associated DRR is $\mathbf{I}_{DRR}^q$, we can define the residual complexity S_{RC}^q of these two images as:

$$\begin{cases} r^q = \mathbf{I}_{X-ray}^q - \mathbf{I}_{DRR}^q \\ c^q = DCT\,(r^q) \\ S_{RC}^q = \sum \log\left(\frac{(c^q)^2}{\alpha} + 1\right) \end{cases} \tag{4}$$

  Where $DCT\,(\cdot)$ is the forward multidimensional discrete cosine transforms [24] and α is a trade-off parameter. For all the experiments presented in this study, we empirically chose $\alpha = 0.05$.
- *Step 3: Triangulation-based Point Reconstruction*. Given the updated 2D locations of the projections of a 3D control point with index i, j, k in the input images, an updated 3D position $L_{ijk}^{nr,t+1}$ of the control point $L_{ijk}^{nr,t}$ can be reconstructed from those updated 2D locations via the same triangulation strategy as shown in Fig. 7.
- *Step 4: 2D–3D Registration*. Given two sets of 3D positions $\{L_{ijk}^{nr,t}\}$ and $\left\{L_{ijk}^{nr,t+1}\right\}$ with known correspondences,

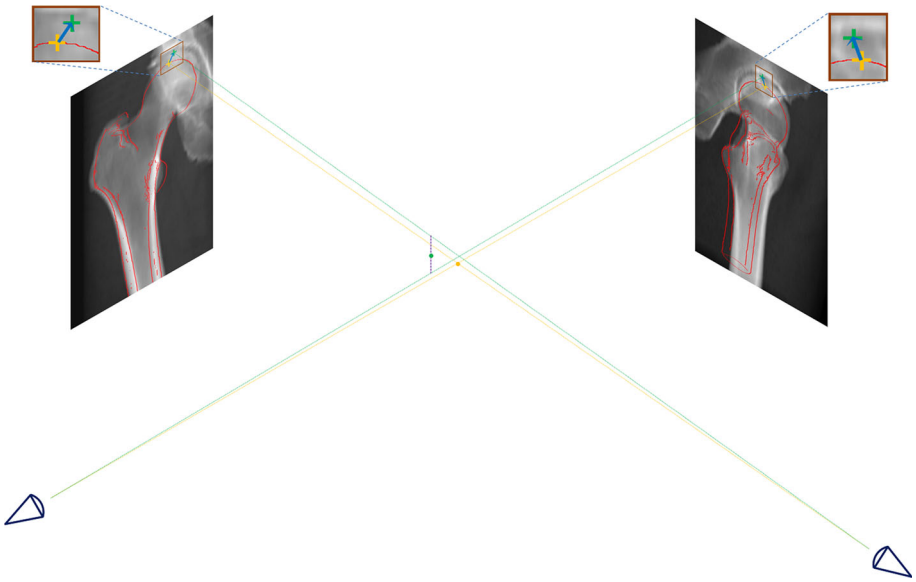

Fig. 7 A schematic view of how to compute the updated position for a control point. The same procedure can be applied to all control points. On both images, we superimposed the contours (*r*ed) extracted from the DRRs for purely visualization purpose. The *yellow dot* is the original location of a control point, and its projection to each X-ray image is shown as *y*ellow cross. An intensity-based 2D–2D image registration is then performed to estimate a 2D transformation to update the projection of this control point to a new location (*g*reen cross in each image). Using a triangulation-based point reconstruction, we can reconstruct a new 3D location of this control point (*g*reen dot) from those updated projection localizations

we compute displacement vectors $\{\overrightarrow{d_{ijk}^{t}} = L_{ijk}^{nr,t+1} - L_{ijk}^{nr,t}\}$ on all control points. We then applied an adaptive regularization on those computed displacement vectors using the method introduced in [25]. After that, we use the 3D tensor product of the familiar 1D cubic B-splines to compute a free-form deformation $\left(T_l^t\right)^{-1}$ at any position of the X-ray reference space [26]. Together with the inverse of the estimated scaled-rigid transformation $\left(T_g\right)^{-1}$, we can use the computed deformation $\left(T_l^t\right)^{-1}$ to warp the template volume $\{\mathbf{I}(\mathbf{x_f})\}$ to the reference space of the X-ray images to obtain an updated volume $\left\{\mathbf{I_c}^{t+1}\right\}$ in the reference space of the given pair of X-ray images.

Repeat *Steps 1–4* a few times (we empirically chose two times) to get an estimation of the local deformation $(T_l)^{-1}$ and a warped volume $\{\mathbf{I_c}\}$.

Experiments and results

Experimental setup and evaluation metrics

After a local institution review board (IRB) approval (Ref.-Nr. KEK-BE: 265/2014), data of 44 patients were used together with data of 49 cadaveric femurs with different shapes to validate the present approach. The 49 cadaveric femurs were divided into two groups. Group I consists of data of 39 cadaveric femurs that were obtained from a previous study [27]. For each femur in this group, we acquired a CT data with a uniform spatial resolution of $0.33 \times 0.33 \times 1.0\,\text{mm}^3$. The mean intensity model created from the CT data of the 39 cadaveric femurs was used as the volumetric template in this study. The mean intensity model was created following the method introduced in our previous work [14] where we introduced a two-stage method to compute an unbiased mean intensity model. Group II contains data from the remaining 10 femurs that were obtained from another previous study [14]. For each femur in this group, we acquired two calibrated images around the proximal femur region using a Siemens Siremobil Iso-C 3D C-Arm (Siemens, Erlangen). All images were eight-bit gray scale, with a size of 768 by 576 pixels. They were calibrated and registered to the associated reference space using the method described in [28]. Additionally, we used anonymized CT data of the 44 patients from a retrospective study investigating medium- and long-term results after unilateral THA. For each patient, CT data of the side without an implantation were used. The intra-slice pixel size of the CT data ranges from 0.55 to 0.83 mm, and their inter-slice spacing ranges from 1.7 to 2.5 mm. We designed and conducted following experiments to validate the present approach.

1. Experiment on 2D grid spacing selection
2. Experiment on 3D grid spacing selection
3. Experiment conducted on C-arm images of the 10 cadaveric femurs
4. Experiment conducted on anonymized data of the 44 patients

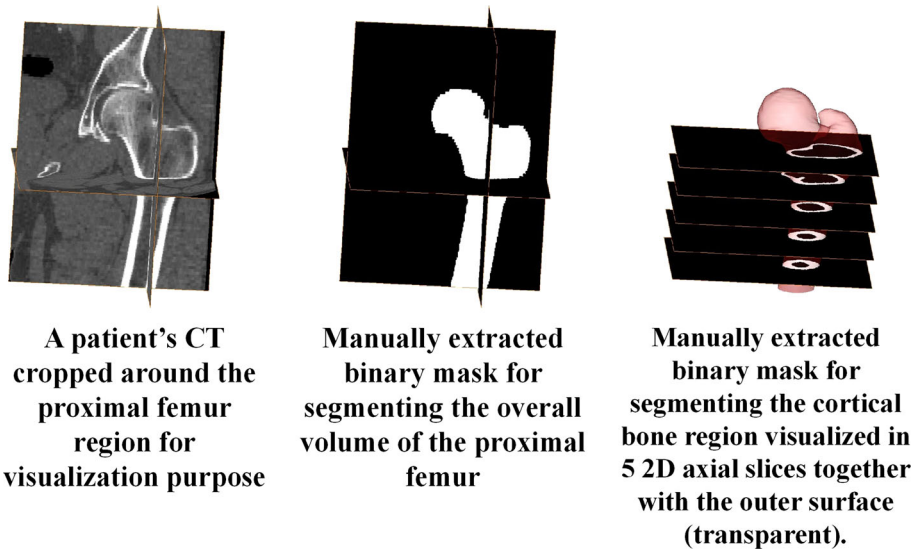

Fig. 8 An example of manual segmentation of a patient's CT, from left to right: the original CT data, the segmented proximal femur volume, and the segmented cortical bone region visualized in five 2D axial slices

In our experiments, we used following metrics.

(a) *2D root mean square (RMS) error*—it is a metric used only in experiment (1) and is defined as the RMS distance from the warped 2D landmarks to the associated, ground truth landmarks.

(b) *3D Dice coefficients (DC)* [29]—it is a metric defined as follows: with L_1 being the binary segmentation obtained from the ground truth CT data and L_2 being the binary segmentation obtained from the reconstructed volume, DC is defined as $\mathrm{DC} = 2\,|L_1 \cap L_2| \,/\, (|L_1| + |L_2|)$. The binary segmentations were manually done with commercial software Amira (www.vsg3d.com/amira). See Fig. 8, middle, for an example.

(c) *3D average surface distance (ASD)* —it is a metric defined as the average distance between vertexes on the surface of a ground truth volume to the outer surface extracted from the binary segmentation of the reconstructed volume.

(d) *3D cortical bone region Dice coefficients (CBRDC)*—it is a metric used in experiments (2) and (4). To calculate such a metric, we first manually segmented the cortical bone region from the ground truth CT data (see Fig. 8, right, for an example), which was regarded as the reference cortical bone region. Furthermore, we manually segment the volumetric template to get a binary mask of the cortical bone region. After the control-point-based 2D–3D registration, we will obtain a deformation field between the volumetric template and the reference space of the 2D X-ray images. We can use the obtained deformation field to warp the binary mask of the cortical bone region of the volumetric template to get the reconstructed cortical bone region, which was then compared with the reference cortical bone region. CBRDC is calculated as the DC between these two binary masks.

(e) *3D cortical bone region ASD (CBRASD)*—it is a metric used in experiments (2) and (4), and is defined as the ASD between the inner cortical bone surfaces extracted from the two binary masks of the segmented cortical bone regions.

Experiment on 2D grid spacing selection

This experiment is conducted on simulated data with the aim to investigate the 2D grid spacing selection on the accuracy of the intensity-based deformable B-spline 2D–2D registration step, which is an important step in determining the accuracy of our 2D–3D reconstruction method. Since the 2D–2D registration at the AP view was independently conducted from the one at the AX view, we have to create two different simulation environments, one for the AP view and the other for the AX view. Below, we focus on the creation of a simulation environment for the AP view, and the same procedure is applied to create a simulation environment for the AX view.

In order to create the training dataset for the automatic contour extraction using the method introduced in [16], we manually annotated the contours of all AP DRRs created from the CT data of the 44 patients. To establish landmark correspondences, we randomly chose one image as the reference and other images are floating images. We evenly sampled a set of 57 landmarks along the contour of the reference image, and the corresponding landmarks in floating images were found by an expectation conditional maximization (ECM)-based deformable shape registration method [30]. In this way, we had established 57 landmark correspondences across all AP DRRs. The same procedure was also applied to the AX DRRs created from the CT data of the 44

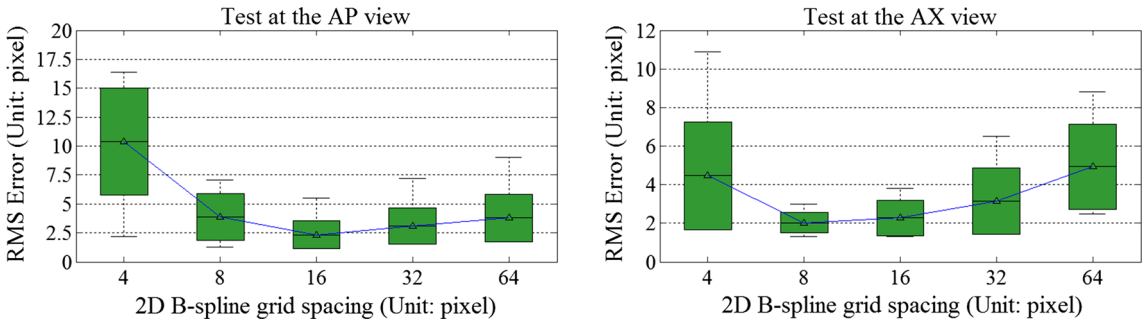

Fig. 9 Results of the experiment on 2D grid spacing selection. *Left* results of the test conducted at the AP view; *right* results for the test conducted at the AX view

patients to establish 54 landmark correspondences across all AX DRRs.

With the known landmark correspondences across all AP DRRs, we found the first 10 pairs of AP DRRs with the biggest RMS distances between the corresponding landmarks, which were then used to create the simulation environment for the AP view. For each chosen pair of AP DRRs, we compute a 2D thin plate spline (TPS) [31] transform based on the corresponding landmarks to warp the femoral region extracted from one DRR as well as the landmarks to the space of the other DRR. The warped images and landmarks together with the original images and landmarks before warping were then used to create a simulation environment to investigate the influence of the 2D grid spacing selection on the accuracy of the intensity-based 2D–2D registration conducted at the AP view. The same procedure was also used to create a simulation environment for the AX view.

For the 2D–2D registrations conducted in both views, we adopted a four-resolution-level image pyramid and a four-level grid spacing pyramid. Thus, the same image pyramid and the same grid spacing pyramid were used in this experiment. Here, the only parameter that we investigated is the size of the final grid spacing. Each time after the intensity-based deformable B-spline 2D–2D registration, we used the obtained deformation field to transform the warped landmarks back to the original space. In this way, we could compute the 2D RMS distance between the warped landmarks and the original landmarks. Figure 9 shows the results when different sizes of 2D grid spacing were used. It was found that for the AP view, a grid spacing of 16 pixels generated the lowest registration error, and for the AX view, a grid spacing of eight pixels generated the lowest error. We thus chose to use these values for other experiments reported in this paper.

Experiment on 3D grid spacing selection

The aim of this experiment was to investigate the 3D grid spacing selection on the 2D–3D reconstruction accuracy. For this purpose, we randomly selected CT data of 10 patients out of the 44 patients. For each selected patient's CT data, two DRRs are generated as the input for this study. As the DRRs were generated from the selected patients' CT data, we can take the patients' CT data as the ground truth to evaluate the 2D–3D registration accuracy when different sizes of 3D grid spacing were used. Each time, in addition to estimating DC and ASD between the ground truth CT data and the reconstructed volume, we further calculated CBRASD and CBRDC to estimate the cortical bone reconstruction accuracy. Figure 10 shows the results when different sizes of 3D grid spacing were used. It was found that a 3D grid spacing of 4 mm generated the best 2D–3D reconstruction results. We thus chose to use a 3D grid spacing of 4 mm for the experiments described below.

Experiment conducted on C-arm images of 10 cadaveric femurs

In this experiment, since there are no surrounding structures around these femurs, for each femur we can directly apply the control-point-based 2D–3D registration algorithm to derive a patient-specific 3D volume from a pair of C-arm images without the requirement to extract the contours first. The reconstruction accuracies were evaluated by randomly digitizing dozens points (see Table 1 for the exact number of points for each case) using a tracked pointer from the surface of each femur and then computing the shortest distances from those digitized points to the associated surface model which was segmented from the reconstructed volume. Table 1 presents for each femur the ASD, which was computed as the average of all shortest distances, and the 95 % percentile error, which was computed after sorting all shortest distances. A mean ASD of 1.29 mm was found, and it took on average four minutes to finish a reconstruction when running on a laptop with 2.5 GHz Intel Core i5 processor and Nvidia GeForce GT 750M graphics card. Figure 11 shows a 2D–3D reconstruction example when the present approach was applied to a pair of calibrated C-arm images.

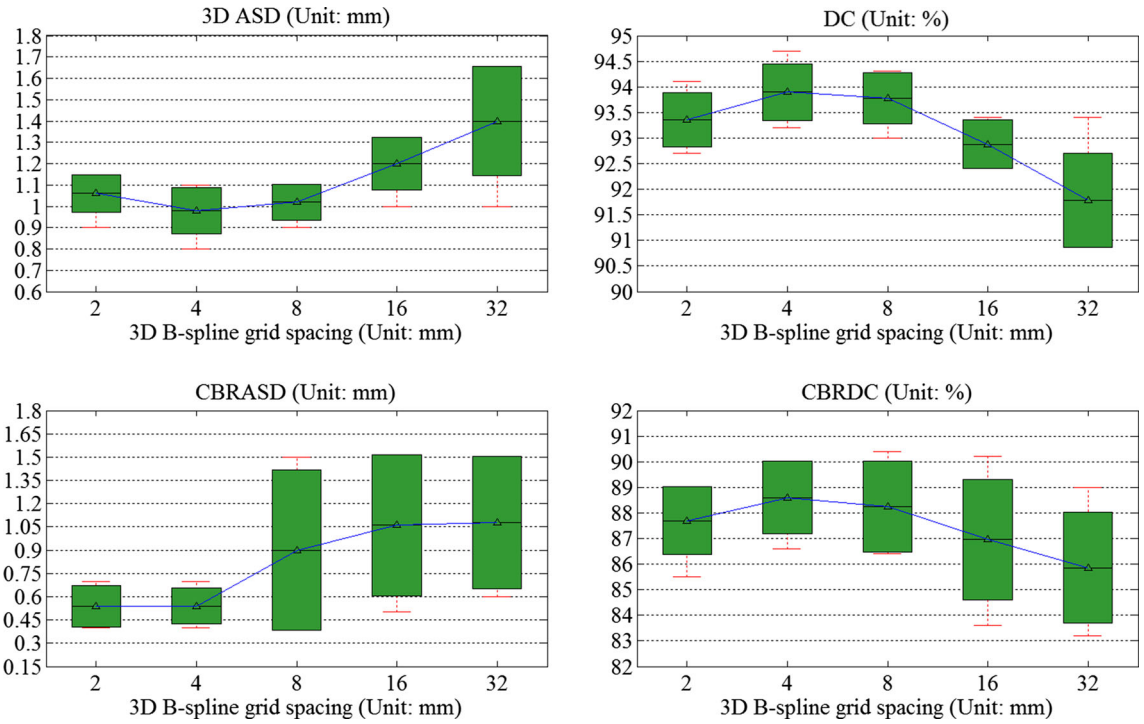

Fig. 10 Results of the experiment on 3D grid spacing selection. *Top-left* 3D ASD results; *top-right* DC results; *bottom-left* CBRASD results; and *bottom-right* CBRDC results

Table 1 Accuracy of reconstructing volumes of 10 cadaveric femurs from calibrated C-arm images

Metrics	#1	#02	#03	#04	#05	#06	#07	#08	#09	#10	Mean
Number of digitized points	96	99	99	99	99	99	99	99	99	99	–
ASD (mm)	1.39	1.03	1.23	1.34	1.14	1.57	1.28	1.43	1.06	1.41	1.29
95 % Percentile error (mm)	3.75	2.13	3.52	3.14	2.41	4.20	2.65	2.99	2.70	3.30	3.08

Experiment conducted on anonymized data of 44 patients

For each patient's CT data, two DRRs are generated as the input for this experiment. Thus, in total we have 44 DRRs generated along each orientation (AP or AX). For DRRs of one orientation, after establishing landmark correspondences across all 44 DRRs, a leave-one-out study was conducted to extract femoral contours using the method introduced in [16]. In addition to estimating DC and ASD between the ground truth CT data and the reconstructed volume, we further calculated CBRASD and CBRDC to estimate the cortical bone reconstruction accuracy. The volume reconstruction errors and the cortical bone region reconstruction errors of the present approach are shown in Table 2. The mean ASD, the mean DC, the mean CBRASD, and the mean CBRDC achieved by the present method are 0.9 ± 0.2 mm, 94.4 ± 1.1 %, 0.7 ± 0.2 mm and 85.1 ± 2.9 %, respectively. Figure 12 shows a comparison of the reconstructed volumes with the associated ground truth volumes. From this figure, one can see that our approach accurately reconstruct the proximal femur morphology and the intramedullary anatomy.

Discussion and conclusion

In this paper, we proposed a new control-point-based 2D–3D volume reconstruction approach and showed its application to a personalized reconstruction of 3D volumes of the proximal femur. With simulated environments, we investigated 2D grid spacing selection. We further investigated the influence of 3D grid spacing selection on the accuracy of the present 2D–3D reconstruction method. The efficacy of the present 2D–3D reconstruction method was finally evaluated on calibrated C-arm images of 10 cadaveric femurs and sim-

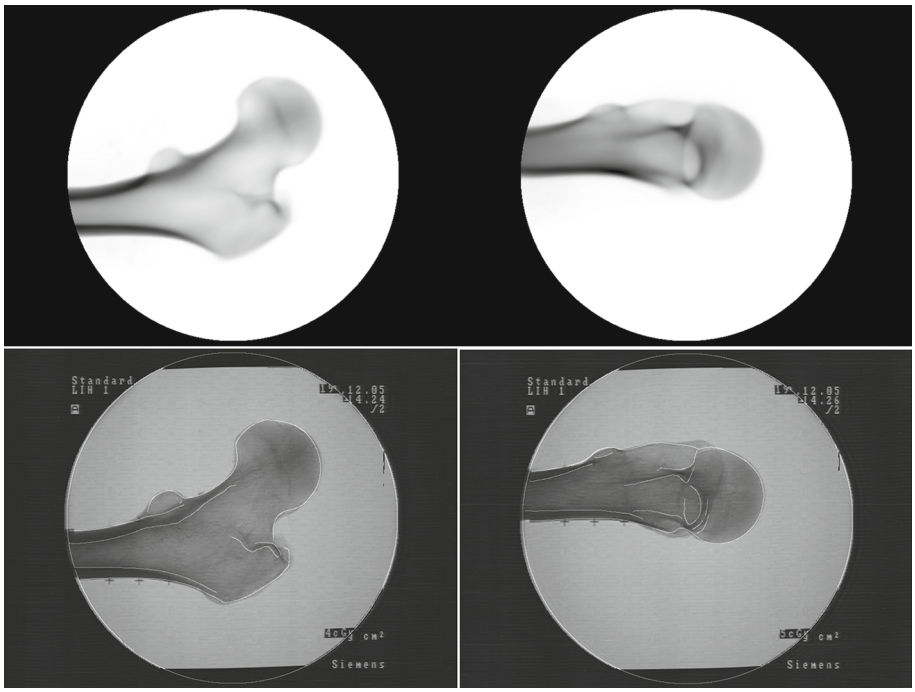

Fig. 11 An example of applying the present 2D–3D reconstruction approach to a pair of C-arm images of a cadaveric femur. *Top row* the DRRs generated from the reconstructed volume; *bottom row* the input C-arm images overlapped with edges extracted from the DRRs shown in the top row

Table 2 Volume and cortical bone region reconstruction accuracy of our approach when evaluated on anonymized data of the 44 patients

Metrics	Mean ± STD	Median	Min	Max
ASD (mm)	0.9 ± 0.2	0.9	0.6	1.4
DC (%)	94.4 ± 1.1	94.5	90.6	95.9
CBRASD (mm)	0.7 ± 0.2	0.7	0.4	1.5
CBRDC (%)	85.1 ± 2.9	85.0	77.2	88.7

ulated X-ray images of 44 patients. Our experimental results showed that our method achieved higher accuracy when evaluated on the simulated X-ray images than on the calibrated C-arm images, suggesting a negative impact of the image calibration error (e.g., Hofstetter et al. [28] reported a mean calibration error of 0.55 ± 0.47 mm when a different C-arm was used) on the reconstruction accuracy.

It is worth to discuss the differences between the present work and other state-of-the-art work on 2D–3D volume reconstruction [11–14]. Unlike previously introduced approaches [11–14] where statistical models are constructed first from a given population of training data and then fitted to the input images, our approach reconstructs a patient-specific 3D volume by fitting one 3D volumetric template to the input images. Although it is difficult to compare the performance of different methods due to the fact that different datasets were used in different studies, both the method introduced in [14] and the present approach have been evaluated on C-arm images of the same 10 cadaveric femurs, which allows us to compare the performance. An average mean reconstruction accuracy of 1.5 mm was reported in [14] while the present approach achieved an average reconstruction accuracy of about 1.3 mm. Although further evaluation is needed before we can draw a definitive conclusion, the evaluation results suggest that the present method is more accurate than the method introduced in [14]. Another advantage of the present method over other state-of-the-art methods [11–14] is that our method is fully automatic and does not need a supervised initialization.

B-spline free-form deformation-based nonrigid registration has been applied to 2D–3D registration tasks before but within different contexts from what has been reported in this study. Shechter et al. [32] reported a 3D method for tracking the coronary arteries through a temporal sequence of biplane X-ray angiography images. Besides the context difference, the work presented in [32] used a different method to find the "best" transformations during motion tracking, i.e., they computed analytical derivatives in 3D and found the solution with a gradient descent method. Another related work was reported in [33] where a triangulation-based 3D point reconstruction scheme was used to facilitate a nonrigid 2D–3D registration of coronary artery models with live fluoroscopy for guidance of cardiac interventions. In order to calculate a reconstructed point for every point in the 3D centerline model, a matching and reconstruction process was used. It was based on a search conducted on the 2D centerlines extracted from X-ray images in the directions that were perpendicular to the projection of the 3D centerline models. The reconstructed points and the associated points in the 3D centerline model were then used to compute a set of 3D displacement vectors for a nonrigid transformation. In contrast, our matching and

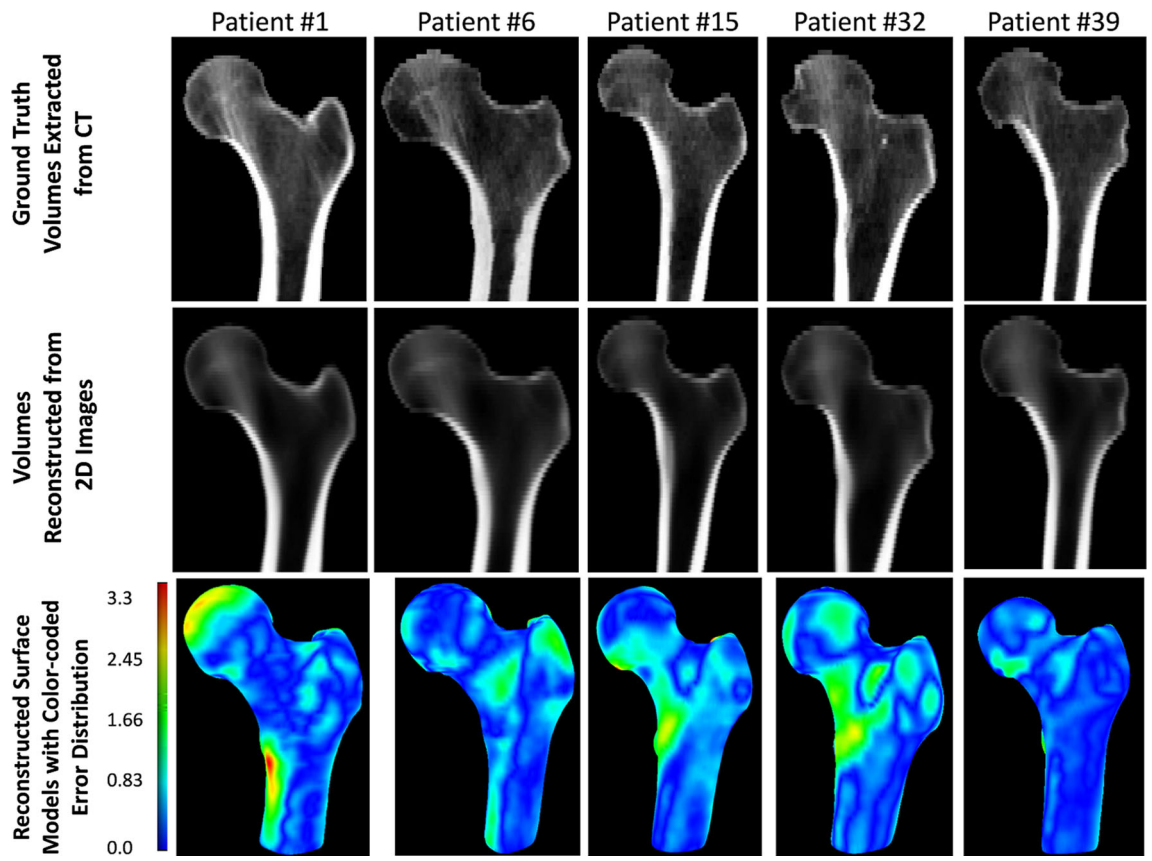

Fig. 12 Comparison of the reconstructed volumes with the associated ground volumes extracted from CT data. *Top two rows* slice views of the ground truth volumes and the reconstructed volumes of five different patients; *bottom row* surface models of the same five patients extracted from the reconstructed volumes with color-coded error distributions

reconstruction process was based on image-based 2D–2D registration.

It is important to note the limitations of the present study. First, due to the usage of a machine-learning-based approach [16] for contour extraction and algorithm initialization, our method requires two calibrated X-ray images acquired from two standard orientations. Despite the fact that we have shown in our previous work [16] that our approach can handle certain deviation from the standard AP orientation, it may fail if the deviation becomes too large. Second, although we conducted comprehensive experiments to evaluate the performance of the present 2D–3D reconstruction method, the method was only evaluated on C-arm images of cadaveric femurs and on simulated X-ray images of patients. Part of our future work will be to investigate its performance on patients' real X-ray images. Nonetheless, the results obtained from the comprehensive experiments demonstrate the efficacy of the present approach. Finally, though for the 2D–3D volume reconstruction experiments reported in this paper we used the mean intensity model created from the CT data of the 39 cadaveric femurs as the volumetric template, in principle our method should work with any volume data of the target anatomy. Then, it is worth to investigate the template selection strategy and its effect on the volume reconstruction accuracy. In this paper, we did not conduct such an investigation, but it is part of our future study.

In summary, we presented a fully automatic approach for personalized reconstruction of 3D volumes of the proximal femur. Comprehensive evaluation demonstrated the efficacy of the present approach. In the future, we will investigate the application of the present methods to other anatomical structures.

Acknowledgments The QCT data of 39 cadaveric femurs in Group I were provided by Dr. P. Zysset. The paper is partially supported by the Japanese-Swiss Science and Technology Cooperation Program and the Swiss National Science Foundation (SNSF) Project No. 205321_138009/1. M. Tannast received support from SNSF via Project No. PP00P3_144856.

Compliance with ethical standards

Conflict of interest The authors have no conflict of interest related to this work.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Sariali E, Mauprivez R, Khiami F, Pascal-Mousselard H, Catonné Y (2012) Accuracy of the preoperative planning for cementless total hip arthroplasty. A randomised comparison between three-dimensional computerised planning and conventional templating. Orthop Traumatol Surg Res 98:151–158
2. Decking R, Puhl W, Simon U, Claes LE (2006) Changes in strain distribution of loaded proximal femora caused by different types of cemetless femoral stem. Clin Biomech 21(5):495–501
3. Aldinger PR, Jung AW, Pritsch M, Breusch S, Thomsen M, Ewerbeck V, Parsch D (2009) Uncemented grit-blasted straight tapered titanium stems in patients younger than fifty-five years of age. Fifteen to twenty-year results. J Bone Joint Surg Am 91(6):1432–1439
4. Hayashi S, Fujishiro T, Hashimoto S, Kanzaki N, Kuroda R, Kurosaka M (2015) The contributing factor of tapered wedge stem alignment during mini-invasive total hip arthroplasty. J Orthop Surg Res 10:52
5. Lecerf G, Fessy MH, Philippot R, Massin P, Giraud F, Flecher X, Girard J, Mertl P, Marchetti E, Stindel E (2009) Femoral offset: anatomical concept, definition, assessment, implications for preoperative templating and hip arthroplasty. Orthop Traumatol Surg Res 95:210–219
6. Huppertz A, Radmer S, Asbach P, Juran R, Schwenke C, Diederichs G, Hamm B, Sparmann M (2011) Computed tomography for preoperative planning in minimal-invasive total hip arthroplasty: Radiation exposure and cost analysis. Eur J Radiol 78:406–413
7. Sodickson A, Baeyens PF, Andriole KP, Prevedello LM, Nawfel RD, Hanson R, Khorasani R (2009) Recurrent CT, cumulative radiation exposure, and associated radiation-induced cancer risks from CT of adults. Radiology 251:175–184
8. Baka N, Kaptein BL, de Bruijne M, van Walsum T, Giphart JE, Niessen WJ, Lelieveldt BP (2011) 2D–3D reconstruction of the distal femur from stereo X-ray imaging using statistical shape models. Med Image Anal 15(6):840–850
9. Le Bras A, Laporte S, Bousson V, Mitton D, De Guise JA, Laredo JD, Skalli W (2004) 3D reconstruction of the proximal femur with low-dose digital stereoradiography. Comput Aided Surg 9:51–57
10. Zheng G, Gollmer S, Schumann S, Dong X, Feilkas T, González Ballester MA (2009) A 2D/3D correspondence building method for reconstruction of a patient-specific 3D bone surface model using point distribution models and calibrated X-ray images. Med Image Anal 13(6):883–899
11. Ahmad O, Ramamurthi K, Wilson KE, Engelke K, Prince RL, Taylor RH (2010) Volumetric DXA (VXA): a new method to extract 3D information from multiple in vivo DXA images. J Bone Miner Res 25(12):2744–2751
12. Sadowsky O, Chintalapani G, Taylor RH (2007) Deformable 2D–3D registration of the pelvis with a limited field of view, using shape statistics. In: Proc. MICCAI 2007, LNCS 4792:519–526
13. Yao J, Taylor RH (2003) Assessing accuracy factors in deformable 2D/3D medical image registration using a statistical pelvis model. In: Proc. ICCV 2003, pp. 1329 – 1334
14. Zheng G (2011) Personalized X-ray reconstruction of the proximal femur via intensity-based non-rigid 2D-3D registration. In: Proc. MICCAI 2011, LNCS 6892:598–606
15. Schumann S, Sato Y, Nakanishi Y, Yokota F, Takao M, Sugano N, Zheng G (2015) Cup implant planning based on 2D/3D radiographic pelvis reconstruction—first clinical results. IEEE Trans Biomed Eng 62:2665–2673
16. Chen C, Zheng G (2014) Fully automatic segmentation of AP pelvis X-ray via random forest regression with efficient feature selection and hierarchical sparse shape composition. Comput Vis Image Underst 126:1–10
17. Chen M, Lu W, Chen Q, Ruchala KJ, Olivera GH (2008) A simple fixed-point approach to invert a deformation field. Med Phys 35:81–88
18. Aprilis G (2013) GPU accelerated volume rendering for use in 2D-3D registration. Master's thesis, University of Bern, Switzerland
19. Klein S, Staring M, Murphy K, Viergever MA, Pluim JP (2010) Elastix: a toolbox for intensity-based medical image registration. IEEE Trans Med Imaging 29(1):196–205
20. Mattes D, Haynor DR, Vesselle H, Lewellen TK, Eubank W (2003) Pet-CT image registration in the chest using free-form deformations. IEEE Trans Med Imaging 22(1):120–128
21. Klein S, Pluim JP, Staring M, Viergever M (2009) Adaptive stochastic gradient descent optimization for image registration. Int J Comput Vision 81(3):227–239
22. Challis JH (1995) A procedure for determining rigid body transformation parameters. J Biomech 28:733–737
23. Myronenko A, Song X (2010) Intensity-based image registration by minimizing residual complexity. IEEE Trans Med Imaging 29(11):1882–1891
24. Strang G (1999) The discrete cosine transform. SIAM Rev 41(1):135–147
25. Myronenko A, Song X (2009) Adaptive Regularization of Ill-posed Problems: Application to Non-rigid Image Registration. CoRR, abs/0906.3323:1-10
26. Rueckert D, Sonoda LI, Hayes C, Hill DL, Leach MO, Hawkes DJ (1999) Non-rigid registration using free-form deformations: application to breast MR images. IEEE Trans Med Imaging 18(8):712–721
27. Dall'Ara E, Luisier B, Schmidt R, Kainberger F, Zysset P, Pahr D (2013) A nonlinear QCT-based finite element model validation study for the human femur tested in two configurations in vitro. Bone 52:27–38
28. Hofstetter R, Slomczykowski M, Sati M, Nolte LP (1999) Fluoroscopy as an imaging means for computer-assisted surgical navigation. Comput Aided Surg 4(2):65–76
29. Dice L (1945) Measures of the amount of ecologic association between species. Ecology 26(3):297–302
30. Zheng G (2013) Expectation conditional maximization-based deformable shape registration. In: CAIP 2013, vol. 1, pp. 548–555
31. Bookstein FL (1989) Principal warps: thin-plate splines and the decomposition of deformations. IEEE Trans Pattern Anal Mach Intell 11(6):567–585
32. Shechter G, Devernay F, Coste-Manière E, Quyyumi A, McVeigh ER (2003) Three-dimensional motion tracking of coronary arteries in biplane cineangiograms. IEEE Trans Med Imaging 22(4):493–503
33. Rivest-Hénault D, Sundar H, Cheriet M (2012) Nonrigid 2D/3D registration of coronary artery models with live fluoroscopy for guidance of cardiac interventions. IEEE Trans Med Imaging 31(8):1557–1572